Xu Zheng

3rd year Ph.D Candidate @ HKUST, Guangzhou — Resident Doctral Researcher @ INSAIT, Sofia

Tel: +86 15942357285, +359 889357285 — Email: zhengxu128@gmail.com — Homepage — Google Scholar — Linkin

Working Experience

- **INSAIT (Institute for Computer Science, Artificial Intelligence and Technology)** — Sofia, Bulgaria
  Resident Doctoral Researcher - Supervisor: Luc Van Gool & Danda Pani Paudel — *Feb. 2025 – Present*

Education

- **The Hong Kong University of Science and Technology (Guangzhou)** — Guangzhou, China
  Ph.D. in Artificial Intelligence - Supervisor: Xuming Hu & Raymond Chi-Wing Wong — *Sep. 2023 – Present*
- **Northeastern University** — Shenyang, China
  B.Eng. in Communication Engineering, M.Eng. in Information and Communication Engineering — *Sep. 2015 – Jun. 2023*

Publications (Citation: 750+)

First Author:

1. **ICCV 2025** (International Conference on Computer Vision). *Reducing unimodal bias in multi-modal semantic segmentation with multi-scale functional entropy regularization.*

2. **ICCV 2025** (International Conference on Computer Vision). *Omnisam: Omnidirectional segment anything model for uda in panoramic semantic segmentation.*

3. **TPAMI 2024** (IEEE Transactions on Pattern Analysis and Machine Intelligence). *360sfuda++: Towards source-free uda for panoramic segmentation by learning reliable category prototypes.*

4. **ECCV 2024** (European Conference on Computer Vision) ***Oral***. *Learning Modality-agnostic Representation for Semantic Segmentation from Any Modalities.*

5. **ECCV 2024** (European Conference on Computer Vision). *Centering the Value of Every Modality: Towards Efficient and Resilient Modality-agnostic Semantic Segmentation.*

6. **CVPR 2024** (IEEE/CVF Conference on Computer Vision and Pattern Recognition). *EventDance: Unsupervised Cross-modal Source-free Adaptation for Event-based Object Recognition.*

7. **CVPR 2024** (IEEE/CVF Conference on Computer Vision and Pattern Recognition). *Semantics, Distortion, and Style Matter: Towards Source-free UDA for Panoramic Segmentation.*

8. **CVPR 2024** (IEEE/CVF Conference on Computer Vision and Pattern Recognition). *UniBind: LLM-Augmented Unified and Balanced Representation Space to Bind Them All.*

9. **ICCV 2023** (IEEE/CVF International Conference on Computer Vision). *Look at the Neighbor: Distortion-aware Unsupervised Domain Adaptation for Panoramic Semantic Segmentation.*

10. **CVPR 2023** (IEEE/CVF Conference on Computer Vision and Pattern Recognition). *Both Style and Distortion Matter: Dual-Path Unsupervised Domain Adaptation for Panoramic Semantic Segmentation.*

11. **ICRA 2024** (IEEE International Conference on Robotics and Automation). *Transformer-cnn cohort: Semi-supervised semantic segmentation by the best of both students.*

12. **PR 2024** (Pattern Recognition, Elsevier). *Distilling Efficient Vision Transformers from CNNs for Semantic Segmentation.*

13. **CBM 2022** (Computers in Biology and Medicine, Elsevier). *Uncertainty-aware deep co-training for semi-supervised medical image segmentation.*

Collaboration / Co-author:

1. **ICCV 2025** (International Conference on Computer Vision). *Unlocking Constraints: Source-Free Occlusion-Aware Seamless Segmentation.*

2. **ICCV 2025** (International Conference on Computer Vision). *CIARD: Cyclic Iterative Adversarial Robustness Distillation.*

3. **ACL 2025 Findings** (Findings of the Association for Computational Linguistics). *MMUnLearner: Reformulating Multimodal Machine Unlearning in the Era of Multimodal Large Language Models.*

4. **ACL 2025 Findings** (Findings of the Association for Computational Linguistics). *A Survey of Mathematical Reasoning in the Era of Multimodal Large Language Model: Benchmark, Method & Challenges.*

5. **ICML 2025** (International Conference on Machine Learning). *RealRAG: Retrieval-augmented Realistic Image Generation via Self-reflective Contrastive Learning.*

6. **CVPR 2025 @ TMM-OpenWorld** (IEEE/CVF Conference on Computer Vision and Pattern Recognition – TMM-OpenWorld Workshop) ***Best Paper Award***. *Benchmarking Multi-modal Semantic Segmentation under Sensor Failures: Missing and Noisy Modality Robustness.*

7. **ECCV 2024** (European Conference on Computer Vision). *EventBind: Learning a Unified Representation to Bind Them All for Event-based Open-world Understanding.*

8. **Pattern Recognition 2024** (Pattern Recognition, Elsevier). *Frozen is Better than Learning: A New Design of Prototype-based Classifier for Semantic Segmentation.*

9. **CAI 2024** (IEEE Conference on Artificial Intelligence). *Interact360: Interactive Identity-driven Text to 360° Panorama Generation.*

10. **ICRA 2024** (IEEE International Conference on Robotics and Automation). *Chasing Day and Night: Towards Robust and Efficient All-day Object Detection Guided by an Event Camera.*

11. **CVPR 2024** (IEEE/CVF Conference on Computer Vision and Pattern Recognition) ***Highlight***. *ExACT: Language-guided Conceptual Reasoning and Uncertainty Estimation for Event-based Action Recognition and More.*

12. **ICCV 2023** (IEEE/CVF International Conference on Computer Vision). *A Good Student is Cooperative and Reliable: CNN-Transformer Collaborative Learning for Semantic Segmentation.*

13. **Computers in Biology and Medicine 2023** (Computers in Biology and Medicine, Elsevier). *Adversarial Co-training for Semantic Segmentation over Medical Images.*

14. **Computers in Biology and Medicine 2022** (Computers in Biology and Medicine, Elsevier). *Uncertainty Teacher with Dense Focal Loss for Semi-supervised Medical Image Segmentation.*

PREPRINTS

First Author:

1. Arxiv 2025. MLLMs are Deeply Affected by Modality Bias.

2. Arxiv 2025. Are Multimodal Large Language Models Ready for Omnidirectional Spatial Reasoning?

3. Arxiv 2025. OmniSAM: Omnidirectional segment anything model for uda in panoramic semantic segmentation.

4. Arxiv 2025. Reducing Unimodal Bias in Multi-Modal Semantic Segmentation with Multi-Scale Functional Entropy Regularization.

5. Arxiv 2025. Retrieval Augmented Generation and Understanding in Vision: A Survey and New Outlook.

6. Arxiv 2024. MAGIC++: Efficient and Resilient Modality-Agnostic Semantic Segmentation via Hierarchical Modality Selection.

7. Arxiv 2024. EventDance++: Language-guided Unsupervised Source-free Cross-modal Adaptation for Event-based Object Recognition.

8. Arxiv 2024. Learning Robust Anymodal Segmentor with Unimodal and Cross-modal Distillation.

9. Arxiv 2024. EIT-1M: One Million EEG-Image-Text Pairs for Human Visual-textual Recognition and More.

10. Arxiv 2023. Deep learning for event-based vision: A comprehensive survey and benchmarks.

Collaboration / Co-author:

1. Arxiv 2025 (***Project Leader***). BiXFormer: A Robust Framework for Maximizing Modality Effectiveness in Multi-Modal Semantic Segmentation.

2. Arxiv 2025. Shifting AI Efficiency From Model-Centric to Data-Centric Compression.

3. Arxiv 2025. Self-Supervised and Generalizable Tokenization for CLIP-Based 3D Understanding.

4. Manifold-aware Representation Learning for Degradation-agnostic Image Restoration.

5. Arxiv 2025 (***Corresponding Author***). EGFormer: Towards Efficient and Generalizable Multimodal Semantic Segmentation

6. Arxiv 2025 (***Corresponding Author***). Towards Omnidirectional Reasoning with 360-R1: A Dataset, Benchmark, and GRPO-based Method.

7. Arxiv 2025 (***Project Leader***). Robust Multimodal Segmentation with Representation Regularization and Hybrid Prototype Distillation.

8. Arxiv 2025 (***Corresponding Author***). Adversarial Robustness for Unified Multi-Modal Encoders via Efficient

Calibration.

9. Arxiv 2025. Split Matching for Inductive Zero-shot Semantic Segmentation.

10. Arxiv 2025. DiMeR: Disentangled Mesh Reconstruction Model.

11. Arxiv 2025 (***Project Leader***). Memorysam: Memorize modalities and semantics with segment anything model 2 for multi-modal semantic segmentation.

12. Arxiv 2025. Unveiling the potential of segment anything model 2 for rgb-thermal semantic segmentation with language guidance.

13. Arxiv 2024 (***Corresponding Author***). Customize Segment Anything Model for Multi-Modal Semantic Segmentation with Mixture of LoRA Experts.

14. Arxiv 2024. Edge Priors Image Inpainting with StyleGAN2.

15. Arxiv 2024 Goodsam++: Bridging domain and capacity gaps via segment anything model for panoramic semantic segmentation.

16. Arxiv 2024. Omnibind: Teach to build unequal-scale modality interaction for omni-bind of all.

17. Arxiv 2024. BrightDreamer: Generic 3D Gaussian Generative Framework for Fast Text-to-3D Synthesis.

18. Arxiv 2024. Image anything: Towards reasoning-coherent and training-free multi-modal image generation.

19. Arxiv 2024. Clip is also a good teacher: A new learning framework for inductive zero-shot semantic segmentation.

20. Arxiv 2022. All one needs to know about priors for deep image restoration and enhancement: A survey.

HONORS AND AWARDS

- Outstanding Master's Thesis in Liaoning Province, China -2023 **(<2 %)**
- Outstanding Graduate of Liaoning Province, China -2023 **(<2 %)**
- National Scholarship of China -2022 **(<2 %)**
- First Class Master Scholarship of Northeastern University - 2020/2021/2022 **(20 %)**
- Outstanding Graduate of Liaoning Province -2019 **(<2 %)**
- Outstanding Graduate of Northeastern University - 2019 **(<2 %)**
- First Class Scholarship of Northeastern University - 2019 **(20 %)**

PROFESSIONAL SERVICES

- Reviewer for journals including:
  International Journal of Computer Vision (IJCV),
  IEEE Transactions on Image Processing (TIP),
  IEEE Transactions on Neural Networks and Learning Systems (TNNLS),
  IEEE Transactions on Multimedia (TMM),
  IEEE Transactions on Circuits and Systems for Video Technology (TCI),
  Neurocomputing,
  Image and Vision Computing (IVC),
  Computers in Biology and Medicine (CBM),
  Machine Vision and Applications (MVA).

- Program Committee / Reviewer for conferences including:
  IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR),
  International Conference on Machine Learning (ICML),
  ACM International Conference on Multimedia (ACM MM),
  European Conference on Computer Vision (ECCV),
  Conference on Neural Information Processing Systems (NeurIPS),
  International Conference on Learning Representations (ICLR), etc.